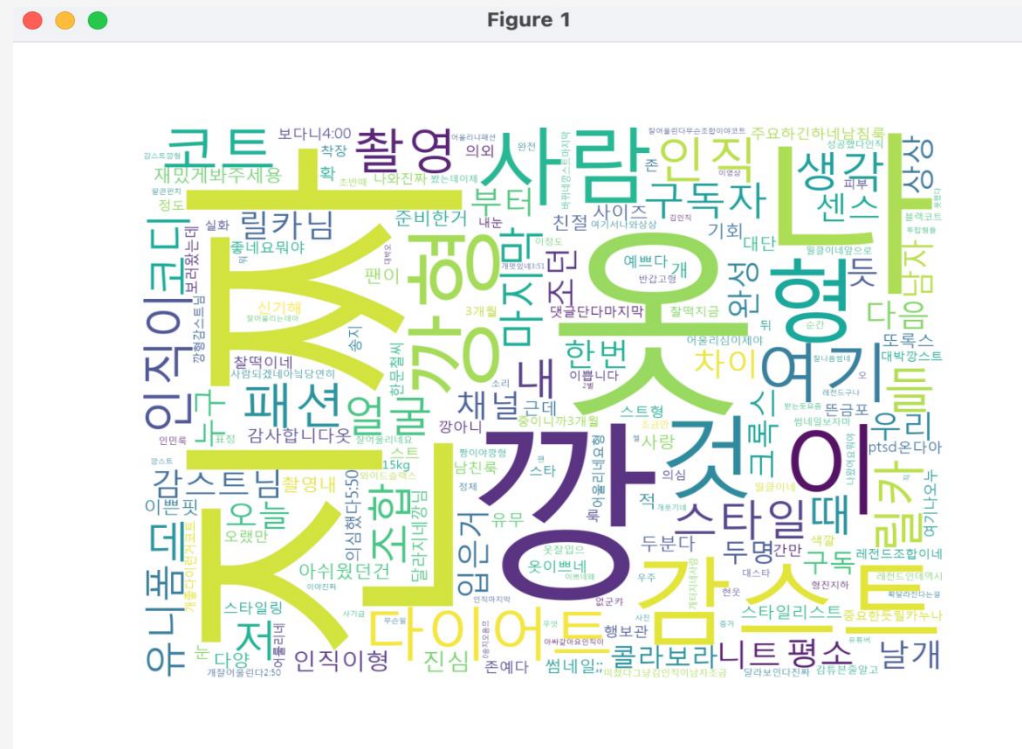


NLP 감정분석 기반
마케팅 시장 분석

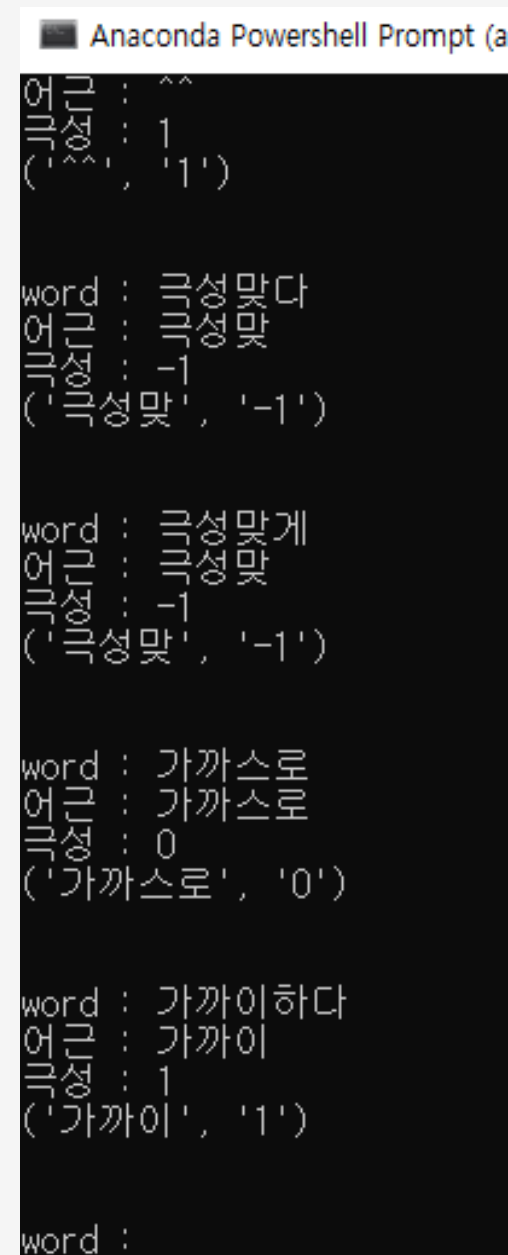
사용자 리뷰 감정분석에 따른 음식점 세부 평가 시스템

빅데이터미네이터

■ YouTube 댓글 감성분석 문제점



B	C	D	E	F	G	H	I
진짜	옷	깡	나	감스트	이	깡형	형
23	11	10	8	8	8	7	7
J	K	L	M	N	O	P	
것	사람	여기	다이어트	코트	촬영	인직	릴카
7	6	6	6	5	4	4	



- 패션 유튜브 영상 댓글 크롤링 및 감성어 추가 test 진행

⇒ 패션 유튜브 댓글 데이터 추출 및 감성어 사전 구축 결과, 대부분의 댓글이 긍정적이거나 혹은 요청의 의미가 담긴 중립적인 내용이 많았습니다. 또한 평가 기준을 세우기 어려운 내용들이 많아 NLP 감정 분석을 진행하기에 적절하지 않은 주제라고 판단했습니다.

- NLP 감정 분석에 적합한 주제에 대해 다시 회의를 진행했습니다. 사용자 감정의 분포, 리뷰 텍스트의 길이, input 데이터의 양, 평가 기준 및 대상 등에 대해 논의를 했고, 최종적으로 “사용자 리뷰 감정분석에 따른 음식점 세부 평가 시스템”으로 주제를 변경했습니다.

“

+

사용자 리뷰 감정분석에 따른 음식점 세부 평가 시스템 😞 😐 😊

해당 주제는 감정 분석을 이용한 요식업 시장의 마케팅 분석 툴 개발을 목표로 하고 있습니다.

단순 수집을 넘어 주요 평가항목에 따른 감정 분석 및 통계를 진행하고
이를 통해 소비자의 요구를 맞추는 마케팅 인사이트 툴 개발을 목적으로 합니다.

이미 정의 되어있는 범용적인 감성어 사전 학습을 기반으로,
더 정확하고 높은 품질의 분석을 위하여 해당 주제에 특화된 추가적인 감성어를 사전에 추가하여
감성 분석에 대한 품질, 정확도를 높이는 것을 목표로 합니다.

01. Input

- Input: 한글로 작성된 음식점 리뷰 게시물
- Input 데이터로 사용할 음식점 리뷰/평가 사이트

→ 망고 플레이트

“오늘은 어디서 뭐 먹지?” 망고플레이트가 주변 맛집, 지역 맛집을 추천해드려요.”

: 리뷰, 사용자수, 수도권 등록 음식점 수가 가장 많은 플랫폼

: 지역별 다양한 음식점에 대한 리뷰 텍스트와 별점 평가 정보를 제공하는 앱/사이트

→ 중구 음식점 리뷰 한정 (프로토타입)

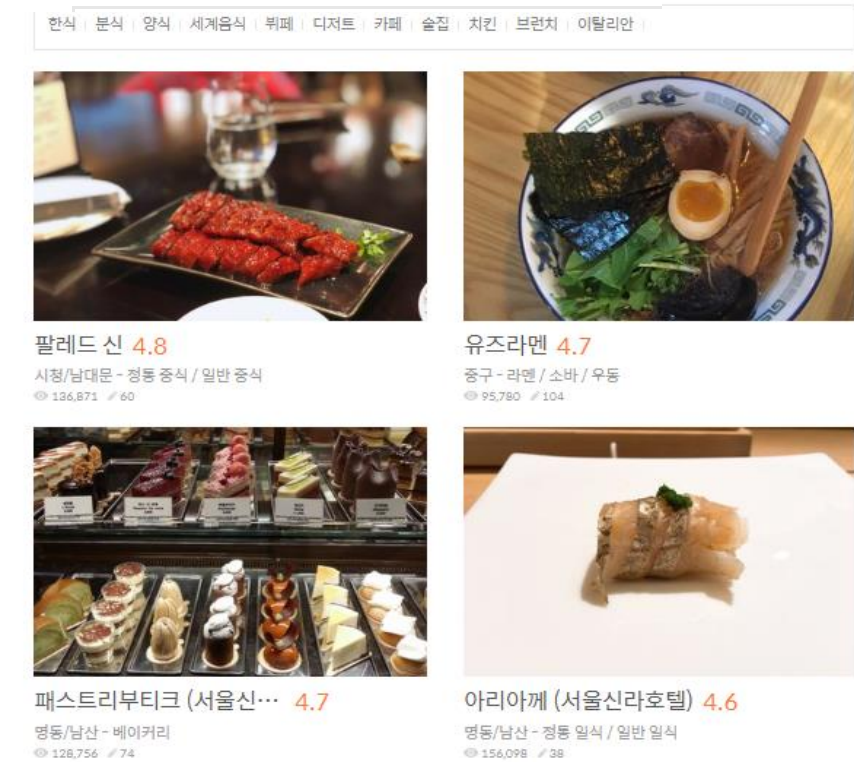
맛집 추천 어플 (사이트) TOP 4

- 4위 – 포잉 (2,151)
- 3위 – 식신 (2,940)
- 2위 – 다이닝 코드 (2,971)
- 1위 – 망고 플레이트 (5,037)

2020. 9. 12.



중구 맛집 인기 검색순위



+

02. 주요 기술

1) 데이터 수집: 크롤링, 전처리

→ 크롤링: 전처리 및 토큰화

→ 이모티콘/특수문자/영어 제거 후 형태소 단위 추출

2) 감정 분석: 평가 항목별 감성어 사전 구축과 감정 분석

→ 범용 감성어 사전: KNU 한국어 감성사전 + 주제 특화 감성 단어 추가

→ 평가 항목:

- | | | |
|-----|--------------------|--------|
| ① 맛 | ②환경(소음, 위치, 분위기 등) | ③서비스 |
| ④가격 | ⑤위생(인테리어, 음식) | ⑥재방문의사 |

3) 결과 도출: 평가 항목별 결과 도출 및 시각화

→ 음식점 이름, 분석한 리뷰 개수, 평가 항목별 점수를 결과로 도출

→ 점수로 도출된 데이터를 그래프 등 시각화된 형태로 제공

Good taste
Bad taste

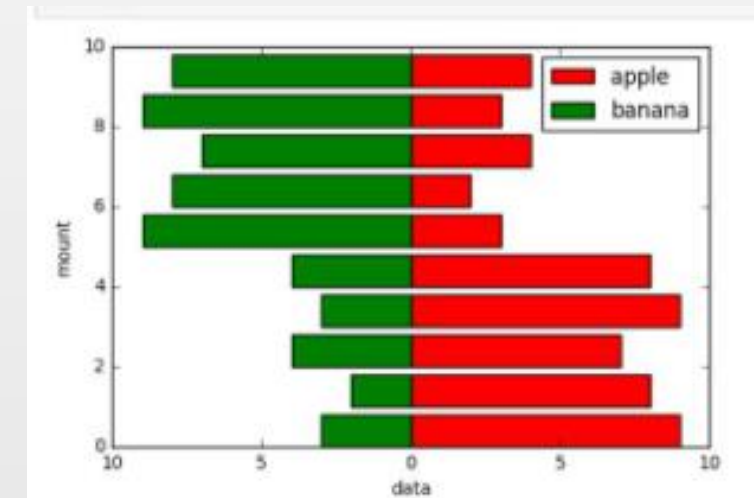
Good Atmosphere
Bad Atmosphere

Good Service
Bad Service

Reasonable price
Unreasonable price

Good Hygiene
Bad Hygiene

Will Revisit
Will not Revisit



● Python

03. Output

→ Output 형태/수준:

항목별 결과를 수치화한 데이터로 제공하는 방식 (프로토타입)

→ Output:

음식점 사용자들의 리뷰를 평가 항목에 따라 감성 분석하여 최종적으로 수치화 한다.

평가 항목: 맛, 환경, 서비스, 가격, 위생, 재방문의사

음식점 이름, 분석에 사용된 리뷰의 개수, 카테고리별 점수를 최종적으로 출력한다.

```
<무학>
총 리뷰 수: 176
-----
Good taste : 140/161 = 0.870
Bad taste : 21/161 = 0.130

Good Atmosphere : 22/26 = 0.846
Bad Atmosphere : 4/26 = 0.514

Good Service : 65/68 = 0.956
Bad Service : 3/68 = 0.004

Reasonable price : 30/47 = 0.638
Unreasonable price : 17/47 = 0.362

Good Hygiene : 60/63 = 0.9523
Bad Hygiene : 3/63 = 0.0476

Will Revisit : 16/20 = 0.8
Will not Revisit : 4/20 = 0.2
-----
```

Ex)



감사합니다 :))

빅데이터미네이터